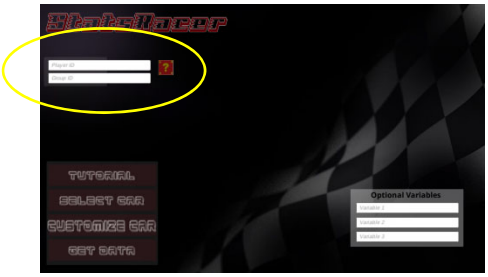
Designing Racer Experiments:

Introduction (Class activity)

In this lab, you will be using an online racing game. In part one of the lab, **your goal is to determine whether the Bayes car or the Gauss car is faster.** In part two, you will **design an experiment to identify car design features to create the "best" car**. Each car you select can be raced multiple times on a track and the time to complete each race will be recorded. After you race the cars, the data will be available for you to analyze.

To play *Racer,* go to the following URL: https://www.stat2games.sites.grinnell.edu/games/racer20.html

- Click to **Run Game** (Note: *This site may take a few minutes to load.*)



Input a **Player ID**. This will be on the internet, so you probably *do not want to use a name that will identify you.*

Record your Player ID:

Player ID: _____

Use the **Group ID** below, which will be identical for every person in the class.

Group ID: **Provided by your Instructor**

- Click on the **Tutorial** button and complete the tutorial
- Go to the following website: http://shiny.grinnell.edu/Racer_Visualizations/
    - Enter the **Group ID** for your class. Make sure that **Level and Track** say **Tutorial**.

Individually complete the tasks below based on the tutorial and include your answers/output with your lab submission:

1. Watch the Racer Data video at https://www.youtube.com/watch?v=isWnFFORQVE. Using appropriate data, create a graph of your class data that addresses the question of which car is faster and copy it below. What observations can you make about your groups' Racer data? What was the response variable you selected and why?

2. Do you believe the data shows evidence of a difference between the speeds of the Bayes and Nightingale cars? How confident are you that one car is faster than the other? Give a numerical representation of your confidence (0% confident to 100% confident).

3. Watch the video entitled "Deception at Duke" at https://www.youtube.com/watch?v=eV9dcAGaVU8. How does a Racer experiment relate to errors discussed in the video? Were there any data errors in the tutorial experiment? If so, how should they be handled?

Part 1: Comparing two cars

Task #1: Designing an Experiment

Design an experiment that outlines the data collection strategy that you will implement to determine **whether the Bayes car or the Gauss car is faster.** Your data collection strategy should allow you to conduct a hypothesis test. In your report for the lab, be sure to address the questions below at a minimum:

4. Establish the hypothesis test (it may help to download the data from the tutorial here to examine all variables available; the visualization options also help).

   What is the explanatory variable?

   What is the response variable?

   What are the units for the data?

   State the null and alternative hypotheses (in words or symbols). Before stating the hypothesis think carefully how you plan to analyze the data that will be collected.

5. Consider confounding variables. In order to evaluate whether the car speeds are truly different, we should take into account possible **confounding variables** (variables that the researcher did not include in the study but that might be connected to both the independent variable and the dependent variable). What are potential confounding variables that might get in the way of determining which car is faster (consider variables shown in the data as well as other factors that are not currently measured)?

6. Develop your protocol (instructions) for the experiment to generate the data. You need clear instructions so you can run the experiment the same every time. Write down the final instructions to collect data for this study. Specifically discuss how you will control for the potential confounding variables as well as the sample size/runs that you will do, and who will drive.

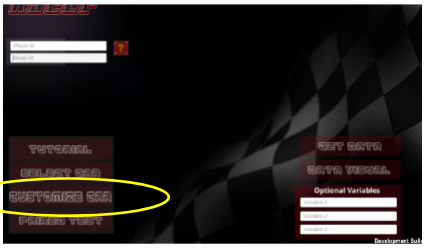
7. Based on your design, what is the appropriate statistical test you will utilize to test the hypotheses?

Task #2: Collect the data

Go to the web site https://www.stat2games.sites.grinnell.edu/games/racer20.html

Establish a Group ID for your experiment by adding 3 letters/numbers to your previous Group ID:
Group ID:

Determine player IDs as appropriate
Player IDs: _____



Conduct the experiment using the protocol decided upon in part 1.

- Click on the **Customize Car** button or the **Select Car** in order to select the appropriate levels for each factor in a given run.

After the experiment is complete, go to the website http://shiny.grinnell.edu/Racer_Visualizations/ in order to download your data.

- Enter the **Group ID** and set **Player ID to all**. Make sure that both **Level** says **ChooseCar**.
- Click the download Racer data button
- Import the data into JMP and save as a JMP data set.
    - Use "File – Open" and select the downloaded .csv file to import to JMP
    - Check each column you will use in the analysis to ensure the data is imported as the appropriate type (right click the variable or column and choose "Column Info" to modify)
    - Use "File – Save" to save as a JMP data table.

Task #3: Analyze the Data

8. Use the "Analyze – Fit Y by X" dialogue to perform complete the hypothesis test for your experiment. Include the following, at a minimum, in your report:

    a. A visualization that helps interpret the test (either from JMP or using the Racer visualization).
    b. The statistical test output.
    c. A summary of the test results and your conclusion.

9. Alternative analysis using randomization

    d. Are there issues with the assumptions of the test you conducted? (Explain)
    e. Perform a randomization test. Include the key output of the test in your report.
    f. How does the randomization test result compare to the parametric test you conducted?

10. Briefly discuss the experiment. What are the limitations of the results? What went well? What improvements/modifications would you make to your protocol were the experiment repeated? What are some future studies/extensions to this experiment?

Prepare notes for group/in class discussion. Some specific questions the class will discuss include which visualization(s) were "best", what the best analysis tools were and discussions of the strength and weaknesses of the experimental design. Be prepared to share your favorite graphs and analysis results.

Part 2: Developing the "Best" Car

Task #4: Considerations for Designing an Experiment that will help you configure the "Best" race car

In this task you will consider choice of design and factors impacting an experiment that considers as many factors as you wish – you will not actually create a complete design or conduct the experiment. The goal is to think about what matters in a more complicated experiment.

11. Select the experimental design you would use based on those discussed in class. Why do you think this design is appropriate?

12. Address the following questions that will impact conducting such an:

    a. How do you define "best"? How does this impact the experiment?

    b. What will be your response variable? Why?

    c. What factors will you include in your experiment? What factors will you hold constant (why)?

    d. What are issues (potential errors, confounding variables) with the data collection that could occur in conducting the experiment (think of at least 2)? Can you address these issues during the experiment – if so how? How will you handle data issues if they still occur regardless of safeguards during the experiment? Are there confounding variables that you cannot control?

    e. What are your resource constraints if any?

    f. If the experiment was run with real cars/tracks are there other issues to consider? Any changes you would make to the experimental design?

Prepare notes for group/in class discussion. Specifically, we will discuss various designs and their potential benefits and issues.

Task #5: Complete the evaluation at: https://grinnell.co1.qualtrics.com/jfe/form/SV_0p4Cz8T1dveoAdf.